基于条件的边界平衡生成对抗网络 *

王硕诚，苟 刚，葛梦园

(贵州大学 计算机科学与技术学院，贵州 550000)

摘 要：生成对抗网络(generative adversarial networks，GAN)是今年来最热门的生成模型之一，使用生成对抗网络和它的一些改进模型可以生成随机图像，或是质量不高的特定图像。目前并没有能够使用简单的网络结构生成高质量特定图像的生成模型，针对这一项任务，提出的方法结合了边界平衡生成对抗网络（boundary equilibrium generative adversarial network，BEGAN）的优点，添加附加条件特征以及均方误差损失，建立条件边界平衡生成对抗网络（conditional-BEGAN，C-BEGAN），使用这种方法提取其中的生成模型用于特定图像的生成，实验结果表明，该方法相比于其他监督类生成模型可以使用更简单的网络达到更快的收敛速度并且能够生成具有更好质量以及多样性的图片。

关键词：生成对抗网络；条件特征；边界平衡；图像生成

中图分类号：TP393 doi: 10.19734/j.issn.1001-3695.2018.10.0842

Conditional boundary equilibrium generative adversarial network

Wang Shuocheng, Gou Gang, Ge Mengyuan

(*College of Computer Science & Technology, Guizhou University, Guizhou* 550000, *China*)

Abstract: Generative adversarial networks (GAN) is one of the most popular generation models of the year. Using the generation antagonistic network and some of its improved models, the model can generate random images, or specific images of low quality. At present, there is no generation model that can use simple network structure to generate high-quality specific images. For this task, the method combines the advantages of boundary equilibrium generative adversarial network(BEGAN) , adds additional condition features and the MSE loss and establishes the conditional boundary equilibrium generative adversarial network(C-BEGAN) . This method used to extract the generation model for specific image generation. Experimental results show that compared with other supervised class generation models, this method can use simpler networks to achieve faster convergence speed and generate images with better quality and diversity.

Key words: generative adversarial network; condition features; boundary equilibrium; image generation

0 引言

图像生成问题一直以来都是一个具有挑战性的问题，对于图像生成模型的建模非常困难，通常通过最大化后验概率进行建模优化，使得在大规模数据下变得极难求解。生成对抗网络（GAN）是 Goodfellow 等[1]在 2014 年提出的一种半监督的生成模型，它是学习数据分布的一种方法，生成对抗网络通过一个生成模型和一个判别模型相互对抗，生成模型拟合数据生成新的样本，判别模型判断新生成的样本和真实样本的真假，达到纳什均衡[2]，由于生成对抗网络无须构造马尔可夫链反复采样，无须人为的构造损失函数等优点，使得 GAN 受到了大量的关注，目前 GAN 已经在图像生成方向得到了广泛的应用。由于生成对抗网络有着非常明显的优点，但同时也存在着收敛困难，梯度消失，梯度爆炸等明显的缺陷，之后对 GAN 进行过多次的改进。深度卷积生成对抗网络（DCGAN）[3]首次将卷积网络引入进生成对抗网络当中，由于卷积网络在图像方面强大的特征提取能力，使得 DCGAN 一度成为图像生成模型的标准。原始的 GAN 由于使用随机的高斯噪声作为输入，只能随机的生成样本，文献[4]通过引入条件模型，使生成对抗网络可以产生特定的样本，实现了数据增强等需求。文献[5]首次将判别器的结构改为自编码器[6]，并提出了一种基于能量的思想，使得 GAN 在高清图片的生成上有了新的进步。文献[7]提出了一种学习分布误差之间的相似度而不是学习分布之间的相似度，目前图像生成模型的主要应用场景都是通过最小化真实数据与生成数据之间的分布来生成图像，如文献[8],通过在条件生成对抗网络中加入重构误差实现了配对数据的图像转换。以及在图像风格转换非常火热的 cyclegan，discogan，dualgan[9~11]，都是在文献[4]的基础上进行了改进，通过引入条件完成不同风格图片之间的转换。另外，利用图像的生成模型，可以完成很多特别的任务，如文献[12,13]。然而，目前所有的监督类生成模型都通过拉近真实数据与生成数据之间的分布来生成图像，利用文献[6]的思想，本文提出了一种基于条件的边界平衡生成对抗网络（C-BEGAN），通过引入条件特征学习分布误差的相似度，并结合判别器为自编码器的特点，在判别器当中添加均方误差损失，使生成模型生成指定的样本，实验结果表明，通过更加简单的网络结构生成的图像，质量和稳定性要高于目前主流的监督模型生成的图像。

1 生成对抗网络原理

传统的 GAN 通过优化生成器，使生成器生成的图片更加接近真实的图片，而通过优化判别器，使生成图片经过判

收稿日期：2018-10-23；**修回日期：**2018-12-14 **基金项目：**国家自然科学基金项目（61562009）

作者简介：王硕诚（1992-），男，辽宁沈阳人，硕士，主要研究方向为机器学习、深度学习、计算机视觉（763793719@qq.com）；苟刚（1979-），男，贵州贵阳人，副教授，硕导，硕士，主要研究方向为模式识别、云计算；葛梦园（1992-），女，河南禹州人，硕士，主要研究方向为流媒体、云计算．

别器后接近于 0，使真实图片经过判别器后接近 1，GAN 的优化函数可以表示为

$$\min_G \max_D V(D,G) = E_{x\sim P_{data(x)}}[\log D(x)] + E_{z\sim P_{z(z)}}[\log(1-D(G(z)))] \quad (1)$$

其中：x 代表真实样本，$P_{data(x)}$ 表示真实样本的分布，z 表示噪声样本，$P_{z(z)}$ 表示噪声分布，通常为高斯分布，Goodfellow 等证明，当生成器与判别器达到最优时，P_z 经过生成器变换后样本的分布与真实图像样本的分布趋于一致。然而原始的 GAN 存在着一些缺点，其中之一就是生成的样本并不受控制，无法产生特定的样本。条件生成对抗网络（CGAN）在 GAN 的基础上添加了条件特征，在生成器中将噪声与条件特征同时作为输入，其目标函数如（2）所示。

$$\min_G \max_D V(D,G) = E_{x\sim P_{data(x)}}[\log D(x\mid c)] + E_{z\sim P_{z(z)}}[\log(1-D(G(z\mid c)))] \quad (2)$$

1.1 条件生成对抗网络

CGAN 的优化过程与 GAN 类似， CGAN 的流程图如图 1 所示。

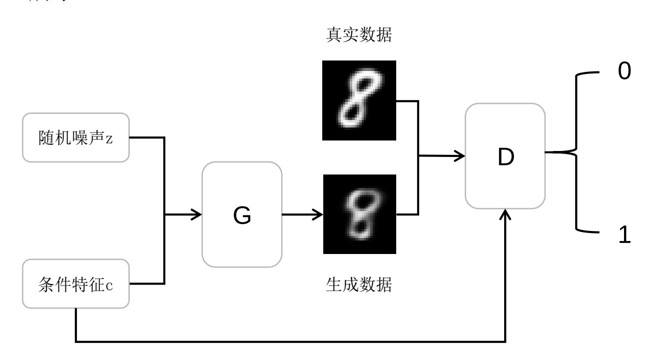


图 1 CGAN 流程图

Fig. 1 CGAN flow chart

CGAN 接收随机噪声与条件特征作为输入，通常条件特征可以为样本的标签数据，同时也可以将广义的图片作为条件特征。CGAN 实现了将无监督的 GAN 改进成了生成特定样本的监督模型。

1.2 基于 Wasserstein 距离生成对抗网络

虽然 GAN 在图像生成上有着出色的效果，然而在训练中却存在着许多的问题，如训练不稳定，梯度消失，梯度爆炸等，因此，大量研究者针对 GAN 的这些缺点进行了改进。WGAN[14]从 GAN 的本质出发进行了改进，并指出 GAN 不稳定的原因是当判别器和生成器都达到最优的时候，优化函数实际变成了优化 JS 散度与 KS 散度问题，并提出了 Wasserstein 距离，Wasserstein 距离的定义如下：

$$W(P_r,P_g) = \inf_{\gamma\sim\prod(P_r,P_g)} E_{(x,y)\sim\gamma}[\|x-y\|] \quad (3)$$

其中：P_r 表示真实数据的分布，P_g 表示生成数据的分布，$\Pi(P_r,P_g)$ 是 P_r 和 P_g 组合起来所有可能的联合分布的集合，对于每一个可能的分布 γ，取得样本 x 与 y，取他们距离的期望值 $E_{(x,y)\sim\gamma}[\|x-y\|]$，在所有可能的分布中取得这个期望值的下界即为 Wasserstein 距离。但由于 $\inf_{\gamma\sim\prod(P_r,P_g)}$ 难以求解，因此根据对偶原理，可将公式转换为

$$W(P_r,P_g) \approx \max_{w:|f_w|_L\le K} \frac{1}{K} E_{x\sim P_r}[f_w(x)] - E_{x\sim P_g}[f_w(x)] \quad (4)$$

其中：$|f_w|_L \le K$ 表示非线性模型 $f_w(x)$ 是 Lipschitz 常数为 K 的函数，在生成对抗网络中，即为判别器结构。之后，通过最小化这一距离来拉近真实样本和生成样本之间的分布。基于 Wasserstein 距离的生成对抗网络彻底解决了模式崩溃与多样性不足等问题。

1.3 边界平衡生成对抗网络

通常生成式对抗网路的判别器都是由一个编码器构成，输入的是一张图片，输出的是这张图片是真实样本的概率。EBGAN 首次将判别器的结构换成自编码器，并提出了一种能量的概念生成更高质量的图片。根据 EBGAN 和 WGAN 的特点，谷歌提出了一种边界平衡生成对抗网络（BEGAN)，通过估计分布误差之间的相似度而不是估计分布之间的相似度来还原分布之间的相似度，若真实数据与生成数据的分布误差相同，那么就认为他们的真实分布相同。BEGAN 的优化函数如（5）所示。

$$\begin{aligned} L_D &= L(x) - k_t L(G(z_D)) \\ L_G &= L(G(z_G)) \\ k_{t+1} &= k_t + \lambda_k(\gamma L(x) - L(G(z_G))) \end{aligned} \quad (5)$$

其中：$L(v) = |v - D(v)|$，判别器是一个自编码器，因此 $L(v)$ 表示图片 v 与经过自编码器 D 产生的图片的 pixel-wise 损失，因此式（5）中 L_D 表明真实图像 x 和 x 经过 D 后的图像的 pixel-wise 损失与生成模型生成的图像 G(z)和经过 D 后的图像 D(G(z))的 pixel-wise 损失之差。当真实图像的误差与生成图像的误差相同时训练完成。判别器不断的最小化 L(x)，最大化 L(G(z))，同时生成器不断最小化 L(G(z))。根据 Wasserstein 距离，在满足条件（6）的前提下有

$$\frac{c_1 + c_2 - 2\sqrt{c_1 c_2}}{\|E(L(x)) - E(L(G(z)))\|^2} \text{ 是一个常数或者是单调递增的} \quad (6)$$

$$W(L(x),L(G(z)))^2 \propto \|E(L(x)) - E(L(G(z)))\|^2 \quad (7)$$

其中：$c_1 c_2$ 是 L(x) 与 L(G(z)) 的方差。理论上当 $E(L(x)) = E(L(G(x)))$ 时是最优的，但此时条件式（6）接近正无穷，因此应用比例控制理论添加超参数 γ，γ 满足：

$$\gamma = \frac{E(L(G(z)))}{E(L(x))} \quad (8)$$

通过对 k 的修改使得判别器的损失达到最优。

2 条件边界平衡生成对抗网络

基于条件的边界平衡生成对抗网络（C-BEGAN）综合 BEGAN 与 CGAN 的优点，通过添加条件特征生成特定的高质量图像。C-BEGAN 的流程如图 2 所示。

像素化损失
真实样本
D
随机噪声z
条件特征c
G
生成样本
像素化损失

图 2 条件边界平衡生成对抗网络流程

Fig. 2 Conditional boundary equilibrium generative adversarial network flow chart

模型的生成器接收随机噪声与条件特征作为输入，生成一张图片，再将生成的图片与原条件特征输入进判别器当中，生成新的图片。同时判别器接收真实的图片与条件特征作为输入生成新的图片。

2.1　模型结构

基于条件的边界平衡生成对抗网络(C-BEGAN)的生成器是一个解码器，输入为噪声和条件特征，输出为一张图片。判别器是一个自编码器，输入一张图片，经编码后与输入的条件特征拼接，再进行解码。为了突出模型的优点，相比于传统图像生成的生成式对抗网络，如 DCGAN，WGAN，BEGAN，文章构建了更加简单的卷积神经网络结构。

C-BEGAN 的生成器的结构如图 3 所示，以 SVHN 数据集为例，该数据集为 32 像素*32 像素*3 通道街景房屋号码图像，首先将一个 62 维的噪声与一个经过编码后的 10 维的条件特征拼接成 72 维的特征作为生成器的输入，经过一个全连接层后转换成(128,8,8)的三维张量，之后经过一个 factor 为 2 的上采样层形成(128,16,16)的张量，之后通过一个卷积核为 3*3，步长为 1 的卷积层(实际实验中可以将这一层的卷积层添加多个)，再经过一个 factor 为 2 的上采样层形成(128,32,32)的张量，之后继续通过一个卷积核为 3*3，步长为 1 的卷积层形成(64,32,32)的张量，最后通过一个卷积核为 3*3，步长为 1 的卷积核形成最终(3,32,32)的输出图片。

C-BEGAN 的判别器的结构如图 4 所示，与传统的 CGAN 不同，首先将输入的三维张量(3,32,32)经过一个 factor 为 2 的下采样层，并通过一个卷积核为 3*3，步长为 1 的卷积层形成(64,16,16)的张量。将该张量转换成 64*16*16 的一维向量，并与经过编码后的条件特征(10 维)拼接成一个一维向量，之后经过一个全连接层形成 32 维特征向量，再经过一个全连接层并转换（Reshape）成(64,16,16)的张量，之后经过一个 factor 为 2 的上采样层形成(64,32,32)的张量，最后经过一个卷积核为 3*3，步长为 1 的卷积层形成最终(3,32,32)的图片。

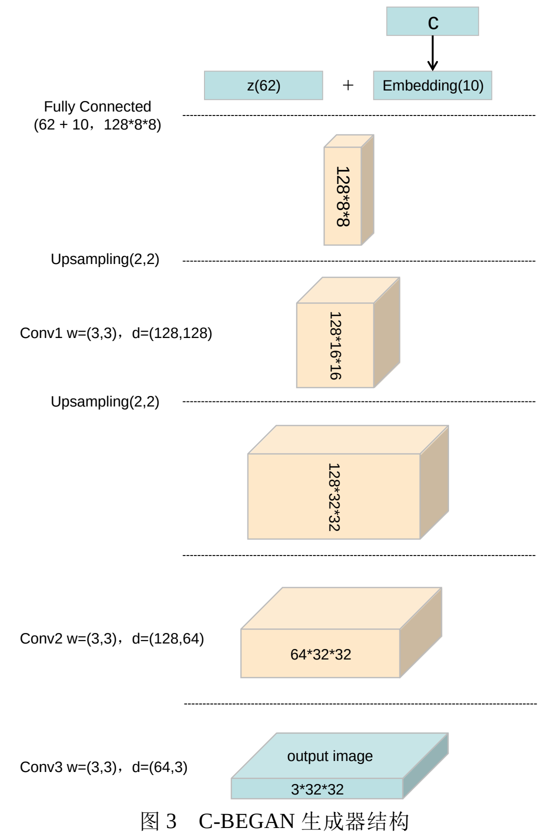


图 3　C-BEGAN 生成器结构

Fig. 3　C-BEGAN generaor structure

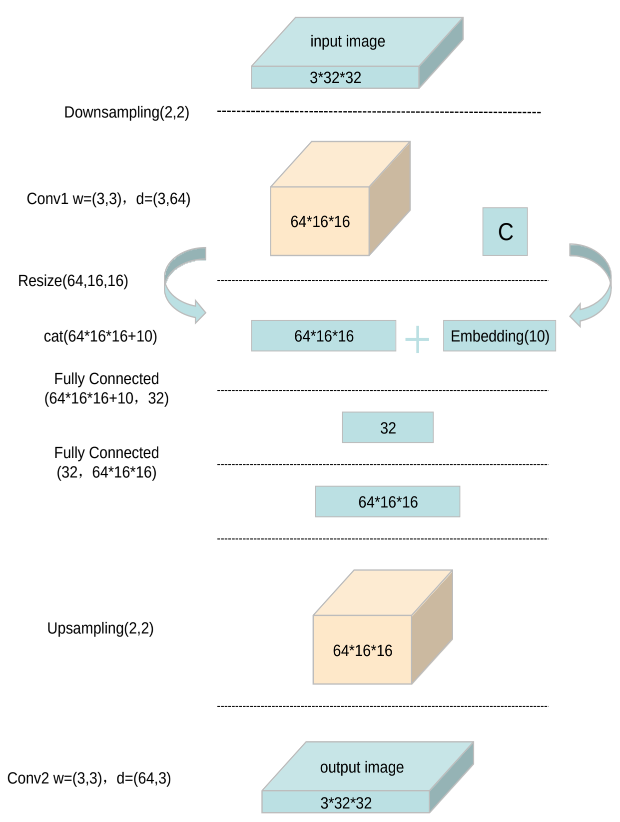


图 4　C-BEGAN 判别器结构

Fig. 4　C-BEGAN discriminator structure

与传统的 CGAN 或 CGAN 的改进版不同，判别器首先并没有将输入图片展开成一维张量再与条件特征进行拼接，而是先对输入图片进行编码，将编码后的张量展开为一维张量再与条件特征进行拼接，这样使得判别器无须将每一层的输出都与条件特征进行拼接也能够生成高质量的图片。并且在较高的分辨率下，直接将图片展开成一维张量会破坏图片的特征结构。

2.2　模型训练

条件边界平衡生成对抗网络的生成器最大化生成图片的 $L(G(z|c))$，判别器一方面最小化真实图片的 $L(x|c)$，一方面最大化生成器所生成图片的 $L(G(z|c))$。由于判别器的结构本质上是一个自编码器，即输入和输出都为图片，为了减小误差，引入均方误差损失，C-BEGAN 的优化函数如（9）所示。

$$\begin{gathered}L_D = L(x|c) - k_t L(G(z_D|c)) \\ L_G = L(G(z_G|c)) + \\ \varphi MSE(D(G(z_G|c)|c), G(z_G|c)) \\ k_{t+1} = k_t + \lambda_k(\gamma L(x|c) - L(G(z_G|c)))\end{gathered} \tag{9}$$

其中生成器需要最小化 $L(G(z_G|c))$，为了提高生成图片的质量，在生成器的后面添加均方误差损失，由于 $L(\bullet)$ 本质为像素化损失，其本身就具有 $L1$ 损失的优点，而 $L1$ 范数对较小误差有着较高的敏感度，MSE 对较大的误差具有较高的敏感度，所以在生成器的后面添加 MSE 损失，使得生成器生成的图片更加的平滑，真实。为了不破坏 Wasserstein 距离的条件，在判别器的后面不加入 MSE 损失。使用超参数 φ 来控制 MSE 损失的权重。每训练一次生成器与判别器后，修改参数 k 的值，使判别器能够达到最优。当生成器与判别器收敛时，生成图片与真实图片分布的误差近乎相等，所以它们的分布近似相等。在训练的过程中，需要对参数 k 做动态的更新，为了保证判别器损失的对称性，需要将 k 的值截取在（0,1）之间，算法使用 Adam 优化器优化损失函数，C-BEGAN 的算法步骤如下所示。

算法 1　C-BEGAN 算法

算法 1 所有的实验设置Batchsize(m=64)，iterations为迭代次数(样本总数与批大小的比值)，epoch为迭代次数（遍历整个数据集的次数）。超参数$\alpha = 0.0002, \varphi = 10, \gamma = 0.75, \lambda_k = 0.01$

for number of training iterations **do**

　for s steps **do**

　　抽样m个服从正态分布的噪声样本 $\{z^{(i)}\}_{i=1}^{m}$

　　抽样m个真实样本数据 $\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

　　生成m个服从正态分布的条件特征 $\{c^{(i)}\}_{i=1}^{m}$

　　计算生成器损失：

　　$L_G = L(G(z_G|c)) + \varphi MSE(D(G(z_G|c)|c, G(z_G|c)))$

　　$g_w \leftarrow \nabla_\omega[\frac{1}{m}|\Sigma_{i=1}^{m} f_\theta(x^{(i)}|c^{(i)}) - \Sigma_{i=1}^{m} x^i| + \frac{1}{m}\Sigma_{i=1}^{m}(f_\theta(x^{(i)}|c^{(i)}) - x^i)^2]$

　　$\omega \leftarrow \omega - \alpha \cdot Adam(\omega, g_\omega)$

　end for

　计算判别器损失：

　$L_D = L(x) - k_\gamma L(G(z_D|c))$

　$f_\theta = \nabla_\theta \frac{1}{m}|\Sigma_{i=1}^{m} f_\theta(x^{(i)}|y^{(i)}) - \Sigma_{i=1}^{m} x^{(i)}| - \frac{1}{m} \cdot k_\gamma \cdot |\Sigma_{i=1}^{m} f_\theta(g_\omega(z^{(i)}|c^{(i)})) - \Sigma_{i=1}^{m} g_\omega(z^{(i)}|c^{(i)})|$

　$\theta \leftarrow \theta - \alpha \cdot Adam(\theta, g_\theta)$

　$k \leftarrow k + \lambda_k(\gamma L_D - L_G)$

　$k = \min(\max(k, 0), 1)$

end for

在实验的过程中，超参数 α，γ，λ_k 都为固定值，而 φ 需要根据不同的数据集而改变，在 MNIST 实验中设为 1，而在 SVHN 实验中设为 10，通常情况下设为 10。在 GAN 的训练过程中，生成器训练的次数通常比判别器训练的次数要多，以防止判别器过于准确而导致生成器无法学习到正确的梯度，基于 Wasserstein 距离的模型不需要多次训练生成器，在实验中将参数 s 设为 1。

3　实验

本文实验在 MNIST 数据集和 SVHN 数据集上进行，实验在 Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz 处理器，一块 NVIDIA Tesla P40 GPU 显卡，Pytorch 环境下进行。

3.1　MNIST 实验

MNIST 数据集是包含 70000 张手写数字的灰度图片，其中每一张图片包含 28×28 个像素点。总共有 0-9 十个类别，其中有 60 000 个训练样本和 10 000 个测试样本。使用数据的时候，由于判别器接收的是一个 32×32 的图像，因此需要将整个数据集转换为 32×32 的张量，其次，使用数据前需要对数据进行归一化处理。在生成器中，除最后一层外所有的卷积层后都接有一个 BatchNorm 层和一个 LeakyReLU 激活层，在最后一层卷积层后面添加一层 Tanh 层。在判别器中，由于需要与条件特征进行拼接，所以在下采样层后只添加一层 ReLU 激活层，不使用 BatchNorm 层[15,16]，而在卷积层后添加了一层 BatchNorm 层和一层 ReLU 激活层。

图 5 为使用 100 个条件特征生成的样本，每行分别输入的条件特征为 0-9，重复 10 次，在实验过程中，仅仅经过了 1 次迭代，模型就学习到了非常准确的特征，并且具有一定的多样性，当到达 50 次迭代后，生成器已经能够生成非常平滑且具有多样性的样本。

图 6 为 C-BEGAN 与目前主流的生成模型加入条件特征后经过 5 次迭代后生成的样本对比图，主要包括有由多层感知机组成的原始条件生成对抗网络（CGAN），深度卷积生成对抗网络（DCGAN）和基于 Wasserstein 距离的 WGAN。除 C-BEGAN 外，其他的生成模型生成的样本均不平滑，存在大量噪声。

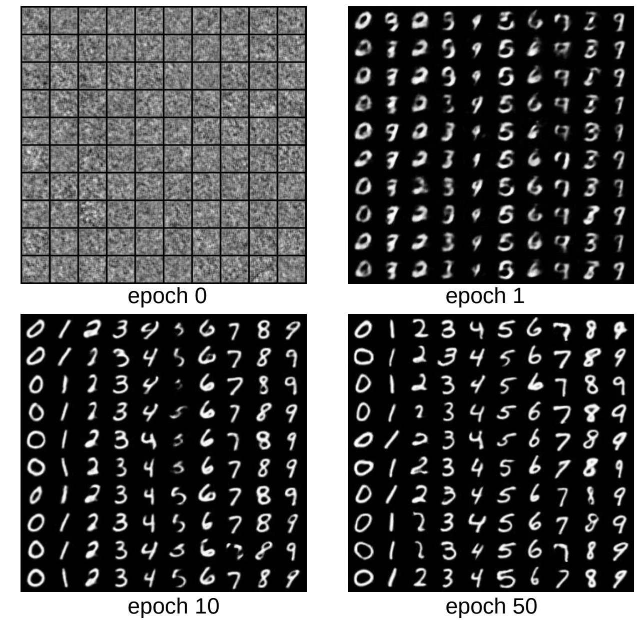


图 5　MNIST 生成样本

Fig. 5　MNIST generate samples

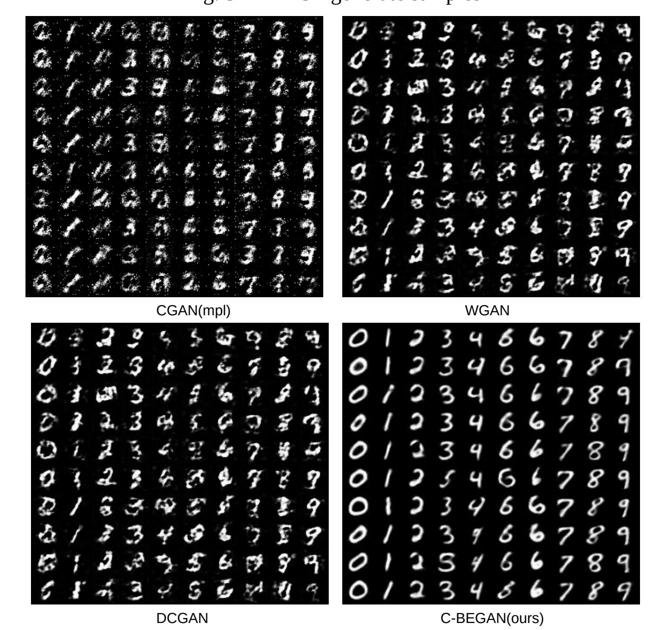


图 6　MNIST 监督模型生成样本对比图

Fig. 6　Contrast figure of MNIST supervised model generate samples

图 7 表示了判别器损失函数的变化趋势，即真实数据与生成数据误差之间的差值。开始的时候损失就呈现了快速下降的趋势，在达到五次迭代之后损失值相对稳定，之后与生成器之间相互对抗，生成的图片变得更加的平滑。

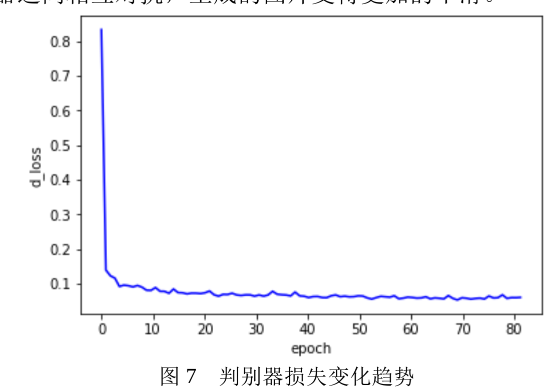


图 7　判别器损失变化趋势

Fig. 7　Discriminator loss trend

图 8 表示了生成器损失函数的变化趋势，即生成的样本与经过自编码器后生成的图片的像素化损失与 *MSE* 损失之和。由于加入了均方误差，损失值较高，但在训练的开始阶段同样下降迅速，在 20 次迭代之后逐渐稳定，此时生成器与判别器的损失平衡。

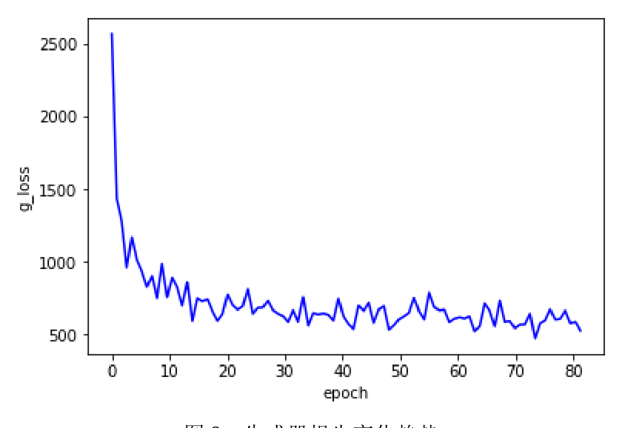


图 8 生成器损失变化趋势

Fig. 8 Generator loss trend

3.2 SVNH 实验

SVHN 数据集是从谷歌街景中查看的房屋号码中收集的，实验使用 Pytorch 提供的共 73 257 张图片，同 MNIST 数据集一样，共 0-9 十个分类，每张图片为 3*32*32 大小的彩色图片，特征量是 MNIST 数据样本的 4 倍。同 MNIST 数据集一样，使用了 BatchNorm 批量归一化层和 LeakyReLU，ReLU 激活层的策略。

图 9 为使用同 MNIST 实验相同的 100 个条件特征生成的样本，经过 10 次迭代后，图像较为平滑，但很多样本并没有学习到条件特征的特点，经过 50 次迭代后，已经能够生成平滑且具有多样性的样本。

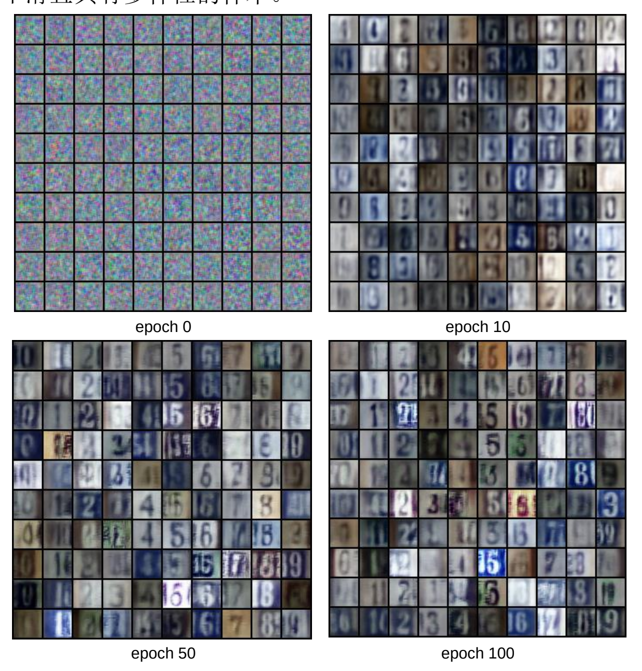


图 9 SVHN 生成样本

Fig. 9 SVHN generate samples

图 10 为不同的生成模型生成的 SVHN 样本对比图，经过 50 次的迭代，其中只有 C-BEGAN 模型生成的样本最为清晰，其他模型生成的样本均较为模糊，其次，除 C-BEGAN 外，其他模型生成样本的数字和背景都不够平滑，因此，C-BEGAN 模型收敛速度更快，生成样本质量更好。



图 10 SVHN 监督模型生成样本对比图

Fig. 10 Contrast figure of SVHN supervised model generate samples

原始的 BEGAN 对模型的依赖性不高，可以达到使用比 DCGAN 更加简单的模型达到预期的效果，但由于本身为无监督模型，但随着训练数据变得复杂，简单的模型会使训练变得困难。如图 11 所示，图中显示了使用上述网络构造同时训练 BEGAN 与 C-BEGAN，图中第一张图片为 BEGAN 训练 20 次迭代后得到的生成样本，第二张图片为 BEGAN 训练 40 次后得到的样本，第四张图片为 C-BEGAN 训练 20 次迭代后得到的样本，可以看到，BEGAN 训练 20 次后得到的样本仅仅学习到了背景特征，而多样性明显不足，产生了模式崩溃，而训练了 40 次迭代后才和 C-BEGAN 训练 20 次迭代后产生质量相近的样本。同时，在用同样的网络训练 BEGAN 的时候，经常会出现训练失败的情况，如第三张图片所示，BEGAN 在训练了 20 次迭代之后依然无法找到正确的梯度。而 C-BEGAN 则能够更好的利用简单的网络进行训练。

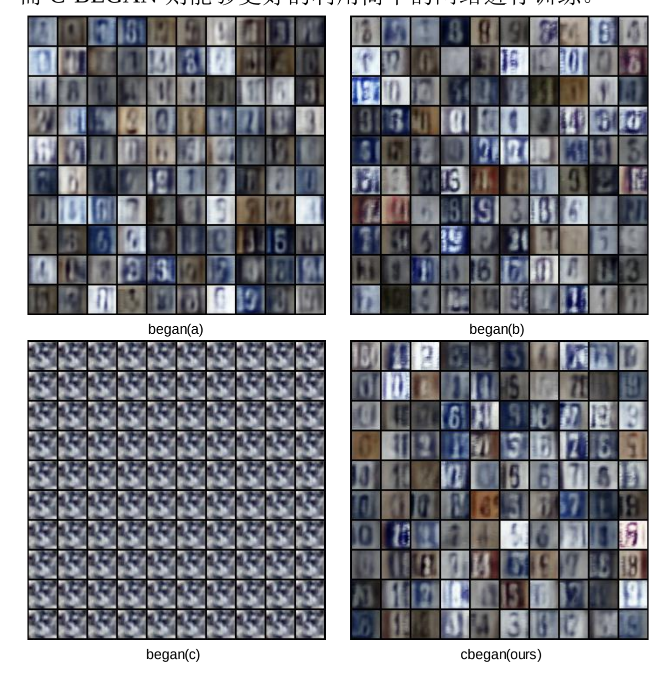


图 11 began 与 cbegan 生成样本

Fig. 11 Generate samples of began and cbegan

为了定量的评估生成图片的质量，本文使用了 Inception Score（IS）评估标准对生成的 1 000 张图片做 10 次平均评估，与其他模型相比较如表 1 所示。

表 1 不同模型 IS 值对比图

Table 1 IS score of different models

使用方法	真实样本	CGAN	DCGAN	WGAN	C-BEGAN(本文算法)
IS 评分	2.38	1.52	1.40	1.48	**1.80**

为了证明这种方法的优良性质，对实验中各种方法的网络参数进行统计，两种实验使用了相同的网络结构，统计结果如表 2 所示。

表 2 不同模型的参数数量对比

Table 2 The number of arguments of different models

网络	CGAN	DCGAN	WGAN	C-BEGAN(本文算法)
生成器	3 944 164	3 341 348	3 343 396	**820 709**
判别器	2 104 421	50 908 004	50 912 100	**1 099 461**

C-BEGAN 生成器参数的数量是其他经典模型参数数量的四分之一左右，判别器的参数是 DCGAN 和 WGAN 判别器参数的五分之一，是原始 CGAN 判别器参数的二分之一，因此，C-BEGAN 网络使用了更简单的网络生成了质量更高对的图片。

4 结束语

目前主流的监督类型的生成对抗网络通过拉近真实数据与生成数据的分布生成样本，实验中发现，加入条件特征后，很多类型的生成对抗网络并不能发挥出原本生成图像的特点，例如从本质上解决了 GAN 模式崩溃等问题的 WGAN，生成特定样本的能力并不好，本文通过使用 BEGAN 拉近数据之间误差的分布，加入条件特征，生成指定的样本，具有非常好的表现，生成器的收敛速度与生成图片的质量和多样性相比于其他模型具有一定的优势。

参考文献：

[1] Goodfellow I J, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets [C]//Proc of International Conference on Neural Information Processing Systems. MIT Press, 2014: 2672-2680.

[2] Ratliff L J, Burden S A, Sastry S S. Characterization and computation of local Nash equilibria in continuous games [C]// Proc of the 51st Annual Allerton Conference on Communication, Control, and Computing. Piscataway, NJ: IEEE Press, 2013: 917-924.

[3] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. arXiv preprint arXiv: 1511. 06434, 2015.

[4] Mirza M, Osindero S. Conditional generative adversarial nets [J]. arXiv preprint arXiv: 1411. 1784, 2014.

[5] Zhao Junbo, Mathieu M, LeCun Y. Energy-based Generative adversarial network [J]. arXiv preprint arXiv: 1609. 03126, 2016.

[6] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. Neural Computation, 2006, 18(7): 1527-1554.

[7] Berthelot D, Schumm T, Metz L. BEGAN: boundary equilibrium generative adversarial networks [J]. arXiv preprint arXiv: 1703. 10717, 2017.

[8] Isola P, Zhu Junyan, Zhou Tinghui, *et al.* Image-to-Image Translation with Conditional Adversarial Networks [J]. arXiv preprint arXiv: 1611. 07004, 2017.

[9] Zhu Junyan, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks [J]. arXiv preprint arXiv: 1703. 10593, 2017.

[10] Kim T, Cha M, Kim H, *et al.* Learning to discover cross-domain relations with generative adversarial networks [J]. arXiv preprint arXiv: 1703. 05192, 2017.

[11] Yi Zili, Zhang Hao, Tan Ping, *et al.* DualGAN: unsupervised dual learning for image-to-image translation [J] arXiv preprint arXiv: 1704. 02510, 2017.

[12] 滕少华，孔棱睿. 基于生成式对抗网络的中文字体风格迁移 [J/OL]. 计算机应用研究，2019，36(11).[2018-0810]. http://www. arocmag. com/article/02-2019-11-055. html. (Teng Shaohua, Kong Lengrui. Chinese fonts style transfer based on generative adversarial networks [J/OL]. Application Research of Computers, 2019, 36(11) . [2018-0810]. http://www. arocmag. com/article/02-2019-11-055. html.)

[13] 梁培俊，刘怡俊. 基于条件生成对抗网络的漫画手绘图上色方法 [J/OL]. 计算机应用研究，2019,36(2). [201801-19]. http://www. arocmag. com/article/02-2019-02-047. html. (Liang Peijuna, Liu Yijuna, Colorization of manga sketch based on conditional generative adversarial networks [J/OL]. Application Research of Computers, 2019, 36(2) . [201801-19]. http://www. arocmag. com/article/02-2019-02-047. html.)

[14] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN [J]. arXiv preprint arXiv: 1701. 07875, 2017.

[15] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//Proc of International Conference on International Conference on Machine Learning. 2015: 448-456.

[16] Simon M, Rodner E, Denzler J. ImageNet pre-trained models with batch normalization [J]. arXiv preprint arXiv: 1612. 01452 2016.